Bipolar disorder is characterised by recurrent hypomanic/manic and depressive episodes that frequently result in hospital admissions. The goal of maintenance treatment is to prevent new mood episodes and relapse. Lithium was the first drug to be used for long-term prophylactic treatment. Although lithium is still widely used, several other drugs have been introduced to prevent mood episodes, such as anticonvulsants (such as valproate, lamotrigine, and carbamazepine) and second-generation antipsychotics (such as olanzapine and quetiapine). Randomised controlled trials (RCTs) have demonstrated the efficacy of several alternatives to lithium, but their results have been criticised and even discounted because of study design and attrition at follow-up. Moreover, RCTs are often conducted in specialised clinics with patients whose illness is more severe and treatment refractory, and they typically employ strict enrolment criteria that limit the generalisability of the findings. For example, bipolar disorder treatment trials might exclude patients with comorbid substance misuse, which is common in clinical practice. Generalisability might be improved by simulating routine clinical practice in randomised effectiveness trials, but such studies are expensive and have difficulty attracting funding. Accordingly, definitive large-scale independent RCTs have not been conducted in psychiatry. The dependence on evidence from industry-supported RCTs has attracted increasingly nihilistic criticism (for example the book Psychiatry Under the Influence) that has arguably reduced public confidence in psychiatry.

For these reasons, convincing naturalistic studies on the real-world effectiveness of psychiatric drugs are important. Observational register studies suggest that lithium is superior to other mood stabilising medications (such as lamotrigine and valproate) in preventing new mood episodes. Such studies have a strong outcome measure (hospital admissions) and are statistically powerful. However, observational studies are limited by the fact that medications are not prescribed randomly. Drugs are selected by the psychiatrist based on the patient’s clinical features, which may be associated with the outcome. This leads to bias called confounding-by-indication. For example, if patients with less severe illness were more likely to be prescribed lithium than valproate, this could create the impression that lithium is more effective than valproate. The aim of this study was to yield a better estimate of the association between drugs used for maintenance treatment in bipolar disorder with rates of psychiatric hospital admissions. We limited confounding-by-indication by a within-individual design, in which all time-stationary covariates, even those that are unobserved, are controlled for by making the individual serve as his or her own control. Thus, we compared periods when patients were medicated with periods when they were not. We studied the effect of lithium, valproate, carbamazepine, lamotrigine, quetiapine and olanzapine on relapse caused by manic, depressive and mixed episodes.

### Method

#### National registries

We used a linkage of several national registries in Sweden: the National Patient Register (NPR), the Swedish Prescribed Drug Register (SPDR), the Swedish national quality assurance register for bipolar disorders (BipolaR), the Total Population Register, Cause of Death Register and Migration Register. The NPR covers in-patient (since 1973) and out-patient (since 2001) psychiatric admissions. The register contains discharge date, main diagnosis and secondary diagnoses based on the International Classifications...
of Diseases (ICD). The in-patient part of the register has full coverage since 1973 and ≥90% of admissions have a registered main diagnosis. The out-patient part was launched in 2001 and the coverage has increased gradually: from 18.2% in 2001 to 87.3% in 2012.12,13 The SPDR contains prescription and dispense dates on all prescribed drugs in Sweden from July 2005. BipolaR, established in Sweden in 2004, contains individualised data on diagnosis, medical interventions and treatment outcomes.15 The Regional Ethics Committee at Karolinska Institutet, Sweden approved the study. The data were anonymised before analysis.

**Patients**

We identified people with bipolar disorder in the NPR according to a modified version of a validated algorithm.16 If the patient had at least two diagnoses of bipolar disorder as either an out-patient or an in-patient in the NPR (ICD-8 codes 1973–1986: 296.0-296.3, 296.8, 296.9; ICD-9 codes 1987–1996: 296A-296E, 296W, 296X; ICD-10 codes from 1997: F30–31),17–19 then this individual was classified as a bipolar disorder ‘case’. Individuals were excluded if they had more than one diagnosis of schizophrenia or schizo-affective disorder, or if the diagnosis of bipolar disorder was only based on ICD-8 296.20 (manic-depressive psychosis, depressed type) and/or ICD-9 296.B (unipolar affective psychosis, melancholic form).16 We also included individuals registered in BipolaR if the algorithm in NPR had not captured them, since diagnoses in BipolaR are considered highly reliable.16 The follow-up time started 1 January 2006 or at the date of first registered diagnosis in the NPR if this occurred after 1 January 2006. For those individuals (n = 262) that were identified only in BipolaR, the start of the follow-up was set to 1 January 2006. Time spent in psychiatric in-patient care was not used as observed time.

**Exposure**

From the SPDR, we extracted data on lithium (anatomical therapeutic chemical (ATC): N05AN01), valproate (ATC: N03AG01), lamotrigine (ATC: N03AX09), carbamazepine (ATC: N03AF01), olanzapine (ATC: N05AH03) and quetiapine (ATC: N05AH04). Treatment periods were defined using sequences of at least two dispense dates. Treatment periods continued until the time between two consecutive dispense dates was longer than 3 months (92 days), then the treatment period ended at the last dispense date. The rationale for using this cut-off is that a single dispensing is, in Sweden, generally limited to a maximum of 3 months’ supply. However, since it is possible to circumvent the rule of 3 months’ supply, and previous research has used other cut-off periods,13,20 we also tested alternative cut-offs (4 and 6 months) in sensitivity analyses. We used SPDR data from 1 July 2005 until 30 June 2010 to determine if the individual was on treatment at the start and end of the study period.

**Outcomes**

The primary outcome in this study was admission to a psychiatric hospital for any reason, defined as at least one overnight stay at a psychiatric clinic. We also used the discharge diagnoses for the secondary outcomes analyses, i.e. admission to hospital because of (a) a manic episode (ICD-10 codes F30-, F310-312), (b) a depressive episode (F313-315, F320-323, and F328-F33), or (c) a mixed episode (F316).19

**Statistical methods**

We defined ‘on’ and ‘off’ periods for all six different medications based on dispense dates, and added psychiatric hospital admission events as outcomes. Our main analysis was the within-individual models where we used stratified Cox regression. Here, each individual entered the model as a separate stratum with an individual baseline hazard function. Different time periods for the same individual, with differing medication status and adjusting for the other medications, were then compared with respect to the outcome, making the individual his/her own control. This design controls for non-time-varying confounders such as gender and genetic makeup, but can still be influenced by time-varying confounders. Only individuals with changing medication status during follow-up contribute information on medication. Other individuals can, however, contribute information on confounders. Additionally, only strata with events contribute information. This statistical design has been used in several previous studies13,20,21 and has been described in detail previously.13 These models were adjusted for previous time spent in psychiatric in-patient care (four-level categorical variable), and age (continuous). The results should be interpreted as the effect of the medication adjusted for both confounders and the five other medications. We also conducted a post-estimation test to compare the effect estimates of the different medications. Since there are 15 pairwise comparisons in each model, we adjusted the P-values using the false discovery rate method.22

To be able to compare our results with previous studies that did not use the within-individual method, we also analysed the risk of hospital admission using a between-individual design. Here, medications were compared between individuals as time-varying covariates. As each individual could potentially contribute with several time-periods of medication, we used robust standard errors. We also conducted a between-individual analysis restricted to those who had received medication at any time during the study period, because untreated individuals might comprise a subset of less ill individuals. These analyses were adjusted for gender, previous time spent in psychiatric in-patient care (four-level categorical variable), and age (continuous) as time-varying covariates. The results should be interpreted as the effect of the medication adjusted for both confounders and the other five medications.

We conducted secondary outcomes analyses (hospital admission because of manic, depressive or mixed episodes) with the within-individual and the two between-individual models. We used SAS 9.3 and R 3.2.2 for all analyses. The texreg package23 was used for table production.

**Sensitivity analyses**

We performed seven sensitivity analyses with the primary outcome to test the robustness of our findings.

(a) We first conducted six separate within-individual analyses, one for each medication, where we excluded all intervals with other medications and only included individuals who at some point received the medication in question. The interpretation of these analyses is straightforward: the effect of the drug is compared with the effect of not being on the drug.

(b) We tested the importance of the sequence of treatments for lithium, which is the first-line treatment for bipolar disorder in Sweden. Here, we only included individuals who had received lithium treatment prior to any of the other study drugs (separate analyses were done for all study drugs except carbamazepine as this drug was rarely used).

(c) We did separate within-individual models for patients that had, and patients that had not, started the observation time on medication to test the influence on starting and stopping medication.
(d) and (e) We used 4 and 6 months as cut-offs defining treatment discontinuation instead of 3 months.

(f) We also tested using 3 months as cut-off but with 30 days added at the end of the medication period. In this analysis, we also included single dispenses of drugs that resulted in a 30-day treatment interval from the dispense date and onwards.

(g) We also analyzed the sample using a 3-day cut-off, instead of 1, as those who are admitted to in-patient care during a weekend may not see a specialist in person until the first working day (Monday). But patients who have been admitted for at least 3 days are quite certain to have been examined by a consultant psychiatrist.

Results

We identified 35,022 individuals with bipolar disorder alive and living in Sweden at some interval between 1 January 2006 and 31 December 2009. Sample characteristics are displayed in Table 1. Of these, 72.3% had a period with any of the six study drugs, with lithium being the most prevalent and carbamazepine the least common. In total, 67.1% of individuals changed their medication status during the study period and a quarter of the study participants were admitted to hospital at least once during follow-up. Online Table DS1 outlines the main discharge diagnoses.

Within-individual analyses

When combining all studied medications into a single variable, medication with any of the study drugs was associated with a reduced rate of admission to a psychiatric hospital (hazard ratios (HR) = 0.67, 95% CI 0.64–0.71). Table 2 shows the results of our main within-individual analyses, adjusting for age, previous time spent in psychiatric in-patient care and the five other drugs. Lithium was associated with a 34% reduction in the rate of admissions to a psychiatric hospital, valproate with 27%, olanzapine with 23%, lamotrigine with 22% and quetiapine with 18% compared with when the individuals were off the respective drug. Unadjusted analyses (not adjusting for age and previous time spent in psychiatric in-patient care) showed similar results as the adjusted analysis for all drugs: lithium (HR = 0.64, 95% CI 0.61–0.69), valproate (HR = 0.72, 95% CI 0.66–0.78), carbamazepine (HR = 0.92, 95% CI 0.77–1.10), lamotrigine (HR = 0.77, 95% CI 0.71–0.82), quetiapine (HR = 0.77, 95% CI 0.71–0.84) and olanzapine (HR = 0.76, 95% CI 0.71–0.82).

Post-estimation head-to-head comparisons of medication effects revealed that lithium was associated with a significantly lower rate of hospital admission than lamotrigine, quetiapine, olanzapine and carbamazepine, while valproate was significantly superior to carbamazepine (Table 3). Lithium, carbamazepine, valproate, quetiapine and olanzapine, but not lamotrigine, were associated with a significantly decreased rate of hospital admission because of a manic episode. Lithium, valproate, lamotrigine, quetiapine and olanzapine, but not carbamazepine, were significantly associated with a reduced rate of admissions because of a depressive episode (Table 2). Finally, lithium and valproate were significantly associated with a reduced rate of admission because of a mixed episode. Head-to-head comparisons of treatment effects on secondary outcomes are given in the online Tables DS2–DS4. These show that lithium, valproate and olanzapine were more effective than lamotrigine in reducing rates of admission to hospital for mania. No other significant differences were found.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Characteristics of the study sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Men (38.4)</td>
</tr>
<tr>
<td>Total participants, n (%)</td>
<td>13,435</td>
</tr>
<tr>
<td>Medication, n (%)a</td>
<td></td>
</tr>
<tr>
<td>Any of the six medications</td>
<td>9,763 (72.7)</td>
</tr>
<tr>
<td>Change in medication status</td>
<td>9,093 (67.7)</td>
</tr>
<tr>
<td>Lithium</td>
<td>6,106 (45.4)</td>
</tr>
<tr>
<td>Valproate</td>
<td>2,158 (16.1)</td>
</tr>
<tr>
<td>Lamotrigine</td>
<td>2,864 (21.3)</td>
</tr>
<tr>
<td>Carbamazepine</td>
<td>525 (3.9)</td>
</tr>
<tr>
<td>Quetiapine</td>
<td>1,459 (10.9)</td>
</tr>
<tr>
<td>Olanzapine</td>
<td>2,624 (19.3)</td>
</tr>
<tr>
<td>Hospital admission during observation time, n (%)</td>
<td>3,419 (25.4)</td>
</tr>
<tr>
<td>Age on 1 January 2006, mean (s.d.)</td>
<td>49.6 (17.0)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Associations between different treatments and admission to psychiatric hospital estimated using within-individual models (n = 35,022)a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
</tr>
<tr>
<td></td>
<td>Manic episodes</td>
</tr>
<tr>
<td></td>
<td>Depressive episodes</td>
</tr>
<tr>
<td></td>
<td>Mixed episodes</td>
</tr>
<tr>
<td>Medication, hazard ratios (95% CI)</td>
<td></td>
</tr>
<tr>
<td>Lithium</td>
<td>0.66 (0.62–0.70)</td>
</tr>
<tr>
<td>Valproate</td>
<td>0.73 (0.67–0.79)</td>
</tr>
<tr>
<td>Carbamazepine</td>
<td>0.92 (0.77–1.10)</td>
</tr>
<tr>
<td>Lamotrigine</td>
<td>0.78 (0.73–0.84)</td>
</tr>
<tr>
<td>Quetiapine</td>
<td>0.82 (0.76–0.89)</td>
</tr>
<tr>
<td>Olanzapine</td>
<td>0.77 (0.72–0.83)</td>
</tr>
<tr>
<td>Events, n</td>
<td>25,330</td>
</tr>
</tbody>
</table>
| a. All models adjusted for previous time spent in psychiatric in-patient care and age.

https://doi.org/10.1192/bjp.bp.116.187989 Published online by Cambridge University Press
Between-individual analyses

When the six medications were combined into one variable, medication was associated with an increased rate of admission to a psychiatric hospital in the between-individual model (HR = 1.06, 95% CI 1.02–1.11). This is likely explained by confounding-by-indication, because when we exclude individuals who never medicated, medication was associated with reduced rates of psychiatric hospital admission (HR = 0.79, 95% CI 0.76–0.82). All between-individual analyses are presented in online Table D55. Note that these analyses are heavily susceptible to confounding-by-indication, even though the counterintuitive positive associations between medication and hospital admissions were attenuated when removing patients that never medicated (online Table D56) during the study period.

Sensitivity analyses

Sensitivity analyses are displayed in online Tables D57–DS13 and generally support the main analysis (online Table D57, Table D59, Table DS11 and Table DS13). However, some differences were notable. If lithium was taken prior to another medication, then the effect of lithium was attenuated (online Table D58). When using 4 months between dispenses as a cut-off for treatment discontinuation, carbamazepine was just significantly associated with reduced rates of hospital admissions (HR = 0.83, 95% CI 0.70–0.99, P = 0.033, online Table D510). Finally, when adding 30 days at the end of each treatment interval and including single dispenses as 30-day treatment intervals, olanzapine and quetiapine were no longer significantly associated with reduced rates of admissions to hospital in the within-individual model (online Table DS12).

Discussion

Our results provide strong evidence that lithium, valproate, lamotrigine, olanzapine and quetiapine lower the risk of psychiatric hospital admission in routine clinical practice. These results are based on the data of 35,022 individuals with bipolar disorder from Swedish national registers, analysed utilising a within-individual model to control for time-stationary confounders. The effects of specific drugs were different and depended on the polarity of the mood episode. Thus, lamotrigine was associated only with a decreased rate of depressive episodes, whereas carbamazepine was solely associated with decreased rate of manic episodes. Lithium, valproate, quetiapine and olanzapine were associated with lower rates of both manic and depressive episodes, supporting their established use as mood stabilisers. Lithium and valproate were the only drugs that were associated with decreased rates of manic, depressive and mixed episodes. These results provide corroborative evidence for the effectiveness of lithium and valproate in preventing mood episodes of any polarity. However, mixed episode was a relatively rare outcome thus lowering the power of finding an effect for the other medications. Moreover, lithium showed a significantly stronger effect than all other medications, except valproate, when we tested the equality of treatment effects in the within-individual analysis of all psychiatric hospital admissions. In line with previous pharmacoepidemiological studies, our between-individual analyses suggest that lithium is more effective than valproate,10 lamotrigine31 and atypical antipsychotics.24 However, these analyses are hampered by confounding-by-indication. For example, our between-individual analysis suggested that quetiapine doubles the rate of manic episodes and that lamotrigine increases the risk of depressive episodes (online Table D55). But this is because quetiapine and lamotrigine are more likely to be prescribed to people prone to mania and depression, respectively. By contrast, our within-individual analyses suggest that quetiapine significantly decreases the risk of manic episodes and lamotrigine decreased the risk of depression. The disagreements between these statistical models clearly demonstrate the importance of controlling for confounding-by-indication in observational pharmacoepidemiological studies. It should be noted, however, that the within-individual design does not control for time-varying confounders, such as, changing disease severity, change of treating physician or changing social circumstances. This precludes direct causal conclusions from this study.

Strengths and limitations

The strengths of this study include the large sample, the use of within-individual analyses to minimise the effect of confounding-by-indication, and the inclusion of several sensitivity analyses to test the robustness of our results. There are also limitations to consider. First, our data on treatment are derived from drug dispenses. We have no data on adherence, which might differ across the studied drugs. However, being a naturalistic study attempting to measure the effectiveness of the study drugs, such differences are embedded into the estimate of the effectiveness of the respective drug. It could thus be argued that this is not a limitation but rather a more correct estimate of the effect of the drug in a real-world setting. Second, our definition on treatment discontinuation (>3 months between two dispense dates) was based on prescriptive standards in Sweden, but it is possible to circumvent these rules. Therefore, we added analyses where we used >4 and >6 months as cut-offs. These sensitivity analyses showed similar results. Third, ascertainment of bipolar disorder cases was based on an algorithm with high positive predictive value but with moderate sensitivity. We might thus have missed
patients with bipolar disorder. We attempted to remedy this by including individuals who had not been diagnosed through the algorithm but were identified in the quality register BipolaR.

Fourth, although our within-individual model handles time-stationary confounders, time-varying confounders are still a potential source of unmeasured confounding. We included age and time spent in hospital as a proxy for illness severity and progression in our model, but there might be other unmeasured confounding. Direct causal interpretation is therefore not possible from our results. Fifth, one form of time-varying confounding is the order in which patients try the respective medication. In a sensitivity analysis where we only included patients who had lithium prior to the other medications, we found that the effect of lithium was attenuated whereas the effect of other medications was slightly enhanced. This suggests that patients who switched from lithium to a second drug were more likely to be non-responders to lithium. Sixth, we were unable to perform separate analyses of treatment effects in bipolar type I and II as there is no information on bipolar subtypes in the NPR. All patients admitted to hospital with mania or mixed episode will, however, necessarily meet criteria for bipolar I disorder. In our sample 39.7% ($n=13,920$) had received an in-patient diagnosis of either mania or mixed episode. Seventh, we did not test all different combinations of pharmacological treatments separately, but instead adjusted for concomitant treatment with other drugs. This is a simplification as the effect of a medication might differ when given together with another drug. Testing all possible combinations of drug treatments was, however, beyond the scope of this study. Eight, predominant polarity is a conceivable effect modifier for the drugs of study. However, the effects of olanzapine and quetiapine were smaller which is striking even though these measures are not completely interchangeable. Thus, the effectiveness of these three drugs in routine clinical practice corresponds well to the effect seen in RCTs.

The relationship with relapse-prevention RCTs

Our effect estimates of lithium, valproate and lamotrigine to prevent mood episodes resemble efficacy estimates in meta-analyses of maintenance treatment in bipolar disorder.2,25,26 In the most recent, the relative risk point estimates for lithium, valproate and lamotrigine were remarkably similar to our hazard ratios,2 and lamotrigine were without enrichment show that effect sizes are larger in enriched samples.27 Finally, it is notable that the lamotrigine maintenance treatment in bipolar disorder.2,25,26 In the most conditions are a study design that avoids as much confounding-by-indication as possible, a defined treatment exposure and meaningful clinical end-points. The results will be notably free of any question of industry or allegiance bias and could improve confidence globally in the practice of psychiatry.

Implications

In conclusion, lithium, valproate, lamotrigine, quetiapine and olanzapine substantially decrease the risk of admissions to hospital in patients with bipolar disorder in a real-world clinical setting, thus largely confirming the findings from relapse-prevention RCTs. This is a reassuring message for both patients and clinicians. However, the effects of olanzapine and quetiapine were smaller compared with results from RCTs, tentatively because of the use of enriched-design RCTs with limited generalisability to the routine clinical setting. High-quality pharmacoepidemiological studies – as well as comparable studies of psychosocial interventions – should assume increasing prominence in developing recommendations for practice in psychiatry. The essential conditions are a study design that avoids as much confounding-by-indication as possible, a defined treatment exposure and meaningful clinical end-points. The results will be notably free of any question of industry or allegiance bias and could improve confidence globally in the practice of psychiatry.

Funding

This study was supported in part by grants from the Swedish Medical Research Council K2014-249-16467/12-51 and K2010-619-21568-01-4, the Swedish Foundation for Strategic Research (KF10-0039), the Swedish Federal Government under the LUA/LF agreement (ALFGBG-142041) and the Sigurd and Elsa Golje Foundation (L2013-0713). Finally, E.J.’s PhD project, which this project is a part of, is partly funded by the University of Gothenburg (L0014/185). The funding sources had no role in the design and conduct of the study; collection, management, analysis, or interpretation of the data; preparation, review, or approval of the manuscript, and decision to submit the manuscript for publication.

Acknowledgements

The authors would like to thank Johan Zetterqvist, MSc and Alexander Viktorin, PhD, for statistical, programming and data management support (both at Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden), Paul Lichtenstein, MD, PhD, for his comments on the manuscript. Finally, this project is conducted as a part of a PhD project, which this project is a part of, is partly funded by the University of Gothenburg (L0014/185). The funding sources had no role in the design and conduct of the study; collection, management, analysis, or interpretation of the data; preparation, review, or approval of the manuscript, and decision to submit the manuscript for publication.

References